Title: NNFWG2

RESULT 7
AAW48713
ID AAW48713 standard; Protein; 2233 AA.
XX
AC AAW48713;
XX
DT 13-OCT-1998 (first entry)
XX
DE HPIV-3 Vero cp45 vaccine L protein.
XX
KW L protein; attenuation; non-segmented; negative sense; vaccine; immunity;
KW single stranded RNA virus; Mononegavirales.
XX
OS Human parainfluenza virus.
XX
PN WO9813501-A2.
XX
PD 02-APR-1998.
XX
PF 19-SEP-1997; 97WO-US16718.
XX
PR 27-SEP-1996; 96US-0026823.
XX
PA (AMCY) AMERICAN CYANAMID CO.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Murphy BR, Randolph VB, Sidhu MS, Tatem JM, Udem SA;
XX
DR WPI; 1998-230710/20.
DR N-PSDB; AAV18274.
XX
PT Recombinantly-generated, attenuated, non-segmented, negative-sense,
PT single stranded RNA virus of order Mononegavirales - having
PT attenuating mutation in 3' genomic promoter region and RNA
PT polymerase gene, useful as vaccine to immunise against such virus
XX
PS Disclosure; Page 283-291; 426pp; English.
XX
CC This sequence represents the Human parainfluenza virus (HPIV-3) type 3
CC vaccine Vero cp45 L protein. This sequence is used in a method which
CC involves the isolation of recombinantly-generated, attenuated,
CC non-segmented, negative-sense, single stranded RNA virus of the order
CC Mononegavirales which have at least 1 attenuating mutation in the 3'
CC genomic promoter region and at least 1 attenuating mutation in the RNA
CC polymerase gene. This RNA virus can be used as a vaccine to immunise an
CC individual against such a virus.
XX
SQ Sequence 2233 AA;

Query Match 100.0%; Score 35; DB 19; Length 2233;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNFWG 5
|||||
Db 1726 NNFWG 1730

Title: NNIWG4

```
RESULT 4
GUNS_ERWCA
ID   GUNS_ERWCA     STANDARD;      PRT;   264 AA.
AC   P16630;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Endoglucanase S precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase S)
DE   (Cellulase S).
GN   CELS.
OS   Erwinia carotovora.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Pectobacterium.
OX   NCBI_TaxID=554;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   STRAIN=SCC3193;
RX   MEDLINE=90337352; PubMed=2379837;
RA   Saarilahti H.T., Henrissat B., Palva E.T.;
RT   "CelS: a novel endoglucanase identified from Erwinia carotovora
RT   subsp. carotovora.";
RL   Gene 90:9-14(1990).
CC   -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
CC       linkages in cellulose, lichenin and cereal beta-D-glucans.
CC   -!- SIMILARITY: BELONGS TO CELLULASE FAMILY H (FAMILY 12 OF GLYCOSYL
CC       HYDROLASES).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M32399; AAA24817.1; -.
DR   PIR; JU0328; JU0328.
DR   InterPro; IPR002594; Glyco_hydro_12.
DR   Pfam; PF01670; Glyco_hydro_12; 1.
DR   ProDom; PD004316; Glyco_hydro_12; 1.
KW   Cellulose degradation; Hydrolase; Glycosidase; Signal.
FT   SIGNAL        1     32
FT   CHAIN        33    264       ENDOGLUCANASE S.
SQ   SEQUENCE   264 AA;  29757 MW;  E6D61388950C77AA CRC64;

  Query Match           97.0%;  Score 32;  DB 1;  Length 264;
  Best Local Similarity  80.0%;  Pred. No. 28;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NNIWG 5
              ||:||
Db         53 NNVWG 57
```

Title: NNLWG1

RESULT 1
US-08-774-065-6
; Sequence 6, Application US/08774065
; Patent No. 5989899
; GENERAL INFORMATION:
; APPLICANT: Bower, Benjamin
; APPLICANT: Clarkson, Kathleen
; APPLICANT: Larenas, Edmund
; APPLICANT: Ward, Michael
; TITLE OF INVENTION: NOVEL OVERSIZED CELLULASE COMPOSITIONS
; TITLE OF INVENTION: FOR USE IN DETERGENT COMPOSITIONS AND
; TITLE OF INVENTION: IN THE TREATMENT OF TEXTILES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENENCOR INTERNATIONAL
; STREET: 925 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/774,065
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Glaister, Debra J.
; REGISTRATION NUMBER: 33,888
; REFERENCE/DOCKET NUMBER: GC368
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-846-7620
; TELEFAX: 415-845-6504
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 77 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-774-065-6

Query Match 100.0%; Score 33; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 NNLWG 5
              |||||
Db         68 NNLWG 72
```

RESULT 2
US-08-774-065-2
; Sequence 2, Application US/08774065
; Patent No. 5989899
; GENERAL INFORMATION:
; APPLICANT: Bower, Benjamin
; APPLICANT: Clarkson, Kathleen
; APPLICANT: Larenas, Edmund
; APPLICANT: Ward, Michael
; TITLE OF INVENTION: NOVEL OVERSIZED CELLULASE COMPOSITIONS
; TITLE OF INVENTION: FOR USE IN DETERGENT COMPOSITIONS AND
; TITLE OF INVENTION: IN THE TREATMENT OF TEXTILES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENENCOR INTERNATIONAL
; STREET: 925 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/774,065
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Glaister, Debra J.
; REGISTRATION NUMBER: 33,888
; REFERENCE/DOCKET NUMBER: GC368
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-846-7620
; TELEFAX: 415-845-6504
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 136 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-774-065-2

Query Match 100.0%; Score 33; DB 2; Length 136;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 NNLWG 5
              |||||
Db         35 NNLWG 39
```

RESULT 3
US-09-216-295-11
; Sequence 11, Application US/09216295
; Patent No. 6268328
; GENERAL INFORMATION:
; APPLICANT: Mitchinson, Colin
; APPLICANT: Wendt, Dan J.
; TITLE OF INVENTION: No. 6268328el Variant EGIII-Like Cellulase Compositions
; FILE REFERENCE: GC555
; CURRENT APPLICATION NUMBER: US/09/216,295
; CURRENT FILING DATE: 1998-12-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Chaetomium brasiliense
US-09-216-295-11

Query Match 100.0%; Score 33; DB 3; Length 194;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 NNLWG 5
              |||||
Db         45 NNLWG 49
```

RESULT 4
US-08-032-848C-10

; Sequence 10, Application US/08032848C
; Patent No. 5475101
; GENERAL INFORMATION:
; APPLICANT: Ward, Michael
; APPLICANT: Clarkson, Kathleen A.
; APPLICANT: Weiss, Geoffrey L.
; APPLICANT: Larenas, Edward
; APPLICANT: Lorch, Jeffrey D.
; TITLE OF INVENTION: Purification and Molecular Cloning of
; TITLE OF INVENTION: EG III Cellulase
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International
; STREET: 180 Kimball Way
; CITY: South San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/032,848C
; FILING DATE: MAR 17 1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Horn, Margaret A.
; REGISTRATION NUMBER: 33,401
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415 742-7356
; TELEFAX: 415 742-7217
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 218 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-032-848C-10

Query Match 100.0%; Score 33; DB 1; Length 218;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 NNLWG 5
              |||||
Db         19 NNLWG 23
```

RESULT 5
US-08-438-870-10
; Sequence 10, Application US/08438870
; Patent No. 5753484
; GENERAL INFORMATION:
; APPLICANT: Ward, Michael
; APPLICANT: Clarkson, Kathleen A.
; APPLICANT: Weiss, Geoffrey L.
; APPLICANT: Larenas, Edward
; APPLICANT: Lorch, Jeffrey D.
; TITLE OF INVENTION: Purification and Molecular Cloning of EG
; TITLE OF INVENTION: III Cellulase
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International
; STREET: 180 Kimball Way
; CITY: South San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94080

```
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/438,870
;       FILING DATE:  May 10, 1995
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Christopher L. Stone
;       REGISTRATION NUMBER:  35,696
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415 742-7555
;       TELEFAX:  415 742-7217
;   INFORMATION FOR SEQ ID NO:  10:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  218 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-438-870-10

  Query Match             100.0%;  Score 33;  DB 1;  Length 218;
  Best Local Similarity   100.0%;  Pred. No. 67;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NNLWG 5
              |||||
Db         19 NNLWG 23


RESULT 6
US-08-169-948B-34
; Sequence 34, Application US/08169948B
; Patent No. 5861271
;  GENERAL INFORMATION:
;     APPLICANT:  Fowler, Timothy
;     APPLICANT:  Ward, Michael
;     APPLICANT:  Clarkson, Kathleen
;     APPLICANT:  Collier, Katherine
;     APPLICANT:  Larenas, Edmund
;     TITLE OF INVENTION:  No. 5861271el Cellulase Enzymes and Systems
;     TITLE OF INVENTION:  For Their Expression
;     NUMBER OF SEQUENCES:  48
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Genencor International
;       STREET:  180 Kimball Way
;       CITY:  South San Francisco
;       STATE:  CA
;       COUNTRY:  USA
;       ZIP:  94080
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/169,948B
;       FILING DATE:  DEC 17 1993
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Horn, Margaret A.
;       REGISTRATION NUMBER:  33,401
;       REFERENCE/DOCKET NUMBER:  GC226
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 742-7536
;       TELEFAX:  (415)742-7217
;   INFORMATION FOR SEQ ID NO:  34:
;     SEQUENCE CHARACTERISTICS:
```

```
;      LENGTH:  218 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-169-948B-34

  Query Match             100.0%;  Score 33;  DB 2;  Length 218;
  Best Local Similarity   100.0%;  Pred. No. 67;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NNLWG 5
              |||||
Db         19 NNLWG 23


RESULT 7
US-08-448-873-34
; Sequence 34, Application US/08448873
; Patent No. 5874276
;  GENERAL INFORMATION:
;    APPLICANT:  Fowler, Timothy
;    APPLICANT:  Ward, Michael
;    APPLICANT:  Clarkson, Kathleen
;    APPLICANT:  Collier, Katherine A.
;    APPLICANT:  Larenas, Edmund
;    TITLE OF INVENTION:  No. 5874276el Cellulase Enzymes and Systems
;    TITLE OF INVENTION:  For Their Expressions
;    NUMBER OF SEQUENCES:  48
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genencor International
;      STREET:  180 Kimball Way
;      CITY:  South San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/448,873
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/169,948
;      FILING DATE:  17-DEC-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stone, Christopher L.
;      REGISTRATION NUMBER:  35,696
;      REFERENCE/DOCKET NUMBER:  GC226D14
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 742-7555
;      TELEFAX:  (415)742-7217
;  INFORMATION FOR SEQ ID NO:  34:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  218 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-448-873-34

  Query Match             100.0%;  Score 33;  DB 2;  Length 218;
  Best Local Similarity   100.0%;  Pred. No. 67;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NNLWG 5
              |||||
Db         19 NNLWG 23
```

RESULT 8
US-08-382-452D-34
; Sequence 34, Application US/08382452D
; Patent No. 6268196
; GENERAL INFORMATION:
; APPLICANT: Fowler, Timothy
; APPLICANT: Clarkson, Kathleen A.
; APPLICANT: Ward, Michael
; APPLICANT: Collier, Katherine D.
; APPLICANT: Larenas, Edmund A.
; TITLE OF INVENTION: NOVEL CELLULOSE ENZYMES AND SYSTEMS
; TITLE OF INVENTION: FOR THEIR EXPRESSION
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International
; STREET: 925 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/382,452D
; FILING DATE: February 1, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Christopher L. Stone
; REGISTRATION NUMBER: 36,696
; REFERENCE/DOCKET NUMBER: GC226-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 742-7555
; TELEFAX: (415)742-7217
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 218 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-382-452D-34

Query Match 100.0%; Score 33; DB 3; Length 218;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNLWG 5
|||||
Db 19 NNLWG 23

RESULT 9
US-09-216-295-1
; Sequence 1, Application US/09216295
; Patent No. 6268328
; GENERAL INFORMATION:
; APPLICANT: Mitchinson, Colin
; APPLICANT: Wendt, Dan J.
; TITLE OF INVENTION: No. 6268328el Variant EGIII-Like Cellulase Compositions
; FILE REFERENCE: GC555
; CURRENT APPLICATION NUMBER: US/09/216,295
; CURRENT FILING DATE: 1998-12-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Trichoderma longibrachiatum

US-09-216-295-1

Query Match 100.0%; Score 33; DB 3; Length 218;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 NNLWG 5
              |||||
Db         19 NNLWG 23
```

RESULT 10
US-08-507-362A-18
; Sequence 18, Application US/08507362A
; Patent No. 6562340
; GENERAL INFORMATION:
; APPLICANT: Bedford, Michael
; Morgan, Andrew
; Fowler, Timothy
; Ward, Michael
; Clarkson, Kathleen
; Collier, Katherine
; Larenas, Edmund
; TITLE OF INVENTION: An Enzyme Feed Additive and Animal Feed Including It
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International
; STREET: 925 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/507,362A
; FILING DATE: 27-Oct-1995
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Castaneda, Janet
; REGISTRATION NUMBER: 33,228
; REFERENCE/DOCKET NUMBER: GC226-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 846-4072
; TELEFAX: (650)845-6504
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 218 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-08-507-362A-18

Query Match 100.0%; Score 33; DB 4; Length 218;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 NNLWG 5
              |||||
Db         19 NNLWG 23
```

RESULT 11
US-08-032-848C-13
; Sequence 13, Application US/08032848C
; Patent No. 5475101

; GENERAL INFORMATION:
; APPLICANT: Ward, Michael
; APPLICANT: Clarkson, Kathleen A.
; APPLICANT: Weiss, Geoffrey L.
; APPLICANT: Larenas, Edward
; APPLICANT: Lorch, Jeffrey D.
; TITLE OF INVENTION: Purification and Molecular Cloning of
; TITLE OF INVENTION: EG III Cellulase
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International
; STREET: 180 Kimball Way
; CITY: South San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/032,848C
; FILING DATE: MAR 17 1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Horn, Margaret A.
; REGISTRATION NUMBER: 33,401
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415 742-7356
; TELEFAX: 415 742-7217
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 221 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-032-848C-13

Query Match 100.0%; Score 33; DB 1; Length 221;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNLWG 5
|||||
Db 21 NNLWG 25

RESULT 12
US-09-146-770-1
; Sequence 1, Application US/09146770
; Patent No. 6187732
; GENERAL INFORMATION:
; APPLICANT: Fowler, Timothy
; TITLE OF INVENTION: Mutant EGIII Cellulase, DNA Encoding
; TITLE OF INVENTION: Such EGIII Compositions and Methods for Obtaining Same
; FILE REFERENCE: GC546
; CURRENT APPLICATION NUMBER: US/09/146,770
; CURRENT FILING DATE: 1998-09-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 232
; TYPE: PRT
; ORGANISM: T. reesei
US-09-146-770-1

Query Match 100.0%; Score 33; DB 3; Length 232;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 NNLWG 5
              |||||
Db         35 NNLWG 39


RESULT 13
US-09-633-084-1
; Sequence 1, Application US/09633084
; Patent No. 6407046
; GENERAL INFORMATION:
;  APPLICANT: Fowler, Timothy
;  TITLE OF INVENTION: Mutant EGIII Cellulase, DNA Encoding
;  TITLE OF INVENTION:  Such EGIII Compositions and Methods for Obtaining Same
;  FILE REFERENCE: GC546
;  CURRENT APPLICATION NUMBER: US/09/633,084
;  CURRENT FILING DATE:  2000-08-04
;  PRIOR APPLICATION NUMBER: 09/146,770
;  PRIOR FILING DATE: 1998-09-03
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 232
;   TYPE: PRT
;   ORGANISM: T. reesei
US-09-633-084-1

  Query Match             100.0%;  Score 33;  DB 4;  Length 232;
  Best Local Similarity   100.0%;  Pred. No. 71;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NNLWG 5
              |||||
Db         35 NNLWG 39

RESULT 3
S12610
cellulase (EC 3.2.1.4) precursor - Aspergillus aculeatus
N;Alternate names: endo-1,4-beta-glucanase
C;Species: Aspergillus aculeatus
C;Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 21-Jul-2000
C;Accession: S12610; S14118; S40186; JQ0458
R;Ooi, T.; Shinmyo, A.; Okada, H.; Murao, S.; Kawaguchi, T.; Arai, M.
Nucleic Acids Res. 18, 5884, 1990
A;Title: Complete nucleotide sequence of a gene coding for Aspergillus aculeatus cellulase
(FI-CMCase).
A;Reference number: S12610; MUID:91016934; PMID:2216782
A;Accession: S12610
A;Molecule type: DNA
A;Residues: 1-237 <OOI1>
A;Cross-references: EMBL:D00546; NID:g217818; PIDN:BAA00435.1; PID:g217819
R;Ooi, T.; Shinmyo, A.; Okada, H.; Hara, S.; Ikenaka, T.; Murao, S.; Arai, M.
Curr. Genet. 18, 217-222, 1990
A;Title: Cloning and sequence analysis of a cDNA for cellulase (FI-CMCase) from Aspergillus
aculeatus.
A;Reference number: S14118; MUID:91064758; PMID:2249253
A;Accession: S14118
A;Molecule type: mRNA
A;Residues: 1-237 <OOI2>
A;Cross-references: EMBL:X52525; NID:g2287; PIDN:CAA36757.1; PID:g2288
A;Accession: S40186
A;Molecule type: protein
A;Residues: 17-18;42-49,'X',51-54,'X';66-79;90-111;136-205,'XX',208-211 <OOI3>
C;Genetics:
A;Introns: 138/2; 212/1
C;Function:
A;Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as cellulose
and lichenin; can hydrolyze such linkages in beta-D-glucans that also contain 1,3-linkages
A;Pathway: cellulose degradation
C;Keywords: glycosidase; hydrolase; polysaccharide degradation; pyroglutamic acid
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-237/Product: cellulase #status experimental <MAT>
```

F;17/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimental

Query Match 100.0%; Score 33; DB 2; Length 237;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 NNLWG 5
              |||||
Db         37 NNLWG 41
```

RESULT 1
GUN_ASPAC
ID GUN_ASPAC STANDARD; PRT; 237 AA.
AC P22669;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Endoglucanase I precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase)
DE (Cellulase) (FI-CMCASE).
OS Aspergillus aculeatus.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5053;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F-50;
RX MEDLINE=91016934; PubMed=2216782;
RA Ooi T., Shinmyo A., Okada H., Murao S., Kawaguchi T., Arai M.;
RT "Complete nucleotide sequence of a gene coding for Aspergillus
RT aculeatus cellulase (FI-CMCase).";
RL Nucleic Acids Res. 18:5884-5884(1990).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=F-50;
RX MEDLINE=91064758; PubMed=2249253;
RA Ooi T., Shinmyo A., Okada H., Hara S., Ikenaka T., Murao S.,
RA Arai M.;
RT "Cloning and sequence analysis of a cDNA for cellulase (FI-CMCase)
RT from Aspergillus aculeatus.";
RL Curr. Genet. 18:217-222(1990).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
CC linkages in cellulose, lichenin and cereal beta-D-glucans.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- INDUCTION: By cellulosic materials and hemicelluloses.
CC -!- MISCELLANEOUS: Will also hydrolyze 1,4-linkages in beta-D-glucans
CC also containing 1,3-linkages.
CC -!- SIMILARITY: BELONGS TO CELLULASE FAMILY H (FAMILY 12 OF GLYCOSYL
CC HYDROLASES).
CC ---

CC ---
DR EMBL; D00546; BAA00435.1; -.
DR EMBL; X52525; CAA36757.1; -.
DR PIR; S12610; S12610.
DR InterPro; IPR002594; Glyco_hydro_12.
DR Pfam; PF01670; Glyco_hydro_12; 1.
DR ProDom; PD004316; Glyco_hydro_12; 1.
KW Cellulose degradation; Hydrolase; Glycosidase; Signal;
KW Pyrrolidone carboxylic acid.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 237 ENDOGLUCANASE I.
FT MOD_RES 17 17 PYRROLIDONE CARBOXYLIC ACID.
SQ SEQUENCE 237 AA; 25560 MW; 8F173571A8AE6931 CRC64;

```
  Query Match             100.0%;  Score 33;  DB 1;  Length 237;
  Best Local Similarity   100.0%;  Pred. No. 18;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NNLWG 5
              |||||
Db         37 NNLWG 41


RESULT 6
O00095
ID   O00095      PRELIMINARY;      PRT;   234 AA.
AC   O00095;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Endo-beta-1,4-glucanase (EC 3.2.1.4).
GN   EGL.
OS   Trichoderma reesei (Hypocrea jecorina).
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Hypocreales; Hypocreaceae; Hypocrea.
OX   NCBI_TaxID=51453;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=QM9414;
RA   Okada H., Tada K., Sekiya T., Yokoyama K., Takahashi A., Tohda H.,
RA   Kumagai H., Morikawa Y.;
RT   "Molecular characterization and heterologous expression of the gene
RT   encoding a low-molec ular-mass endoglucanase from Trichoderma reesei
RT   QM9414.";
RL   Appl. Environ. Microbiol. 64:55-563(1998).
DR   EMBL; AB003694; BAA20140.1; -.
DR   InterPro; IPR002594; Glyco_hydro_12.
DR   Pfam; PF01670; Glyco_hydro_12; 1.
DR   ProDom; PD004316; Glyco_hydro_12; 1.
KW   Glycosidase; Hydrolase.
SQ   SEQUENCE   234 AA;  25159 MW;  DF476EEDE384ADD1 CRC64;

  Query Match             100.0%;  Score 33;  DB 3;  Length 234;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NNLWG 5
              |||||
Db         35 NNLWG 39


RESULT 7
Q8NJY5
ID   Q8NJY5      PRELIMINARY;      PRT;   234 AA.
AC   Q8NJY5;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Endoglucanase.
GN   CEL12A.
OS   Hypocrea koningii.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Hypocreales; Hypocreaceae; Hypocrea.
OX   NCBI_TaxID=97093;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22067395; PubMed=12073090;
RA   Goedegebuur F., Fowler T., Phillips J., van der Kley P.,
RA   van Solingen P., Dankmeyer L., Power S.D.;
RT   "Cloning and relational analysis of 15 novel fungal endoglucanases
RT   from family 12 glycosyl hydrolase.";
RL   Curr. Genet. 41:89-98(2002).
DR   EMBL; AF435069; AAM77712.1; -.
DR   InterPro; IPR002594; Glyco_hydro_12.
DR   Pfam; PF01670; Glyco_hydro_12; 1.
```

```
DR   ProDom; PD004316; Glyco_hydro_12; 1.
SQ   SEQUENCE   234 AA;  25299 MW;  4AFD8486C29FFC44 CRC64;

  Query Match             100.0%;  Score 33;  DB 3;  Length 234;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NNLWG 5
              |||||
Db         35 NNLWG 39


RESULT 8
Q8NJY2
ID   Q8NJY2      PRELIMINARY;      PRT;   237 AA.
AC   Q8NJY2;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Endoglucanase.
GN   CEL12B.
OS   Aspergillus awamori (var. kawachi).
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC   Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX   NCBI_TaxID=40384;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22067395; PubMed=12073090;
RA   Goedegebuur F., Fowler T., Phillips J., van der Kley P.,
RA   van Solingen P., Dankmeyer L., Power S.D.;
RT   "Cloning and relational analysis of 15 novel fungal endoglucanases
RT   from family 12 glycosyl hydrolase.";
RL   Curr. Genet. 41:89-98(2002).
DR   EMBL; AF435072; AAM77715.1; -.
DR   InterPro; IPR002594; Glyco_hydro_12.
DR   Pfam; PF01670; Glyco_hydro_12; 1.
DR   ProDom; PD004316; Glyco_hydro_12; 1.
SQ   SEQUENCE   237 AA;  25710 MW;  4DBDC8563E7CD021 CRC64;

  Query Match             100.0%;  Score 33;  DB 3;  Length 237;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NNLWG 5
              |||||
Db         37 NNLWG 41


RESULT 9
O13454
ID   O13454      PRELIMINARY;      PRT;   239 AA.
AC   O13454;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Endo-1,4-beta-glucanase (EC 3.2.1.4).
GN   CELA.
OS   Aspergillus oryzae.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC   Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX   NCBI_TaxID=5062;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KBN616;
RX   MEDLINE=97161783; PubMed=9008887;
RA   Kitamoto N., Go M., Shibayama T., Kimura T., Kito Y., Ohmiya K.,
RA   Tsukagoshi N.;
RT   "Molecular cloning, purification and characterization of two endo-1,4-
RT   beta-glucanases from Aspergillus oryzae KBN616.";
RL   Appl. Microbiol. Biotechnol. 46:538-544(1996).
DR   EMBL; D83731; BAA22588.1; -.
```

DR InterPro; IPR002594; Glyco_hydro_12.
DR Pfam; PF01670; Glyco_hydro_12; 1.
DR ProDom; PD004316; Glyco_hydro_12; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 239 AA; 26096 MW; C0F850E5DFEB455D CRC64;

Query Match 100.0%; Score 33; DB 3; Length 239;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 NNLWG 5
              |||||
Db         35 NNLWG 39
```